



Schaeffer, J., Benton, M. J., Rayfield, E. J., & Stubbs, T. L. (2019). Morphological disparity in theropod jaws: comparing discrete characters and geometric morphometrics. *Palaeontology*, 63(2), 283-299. <https://doi.org/10.1111/pala.12455>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1111/pala.12455](https://doi.org/10.1111/pala.12455)

[Link to publication record in Explore Bristol Research](#)
PDF-document



This is the final published version of the article (version of record). It first appeared online via Wiley at <https://onlinelibrary.wiley.com/doi/10.1111/pala.12455> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

MORPHOLOGICAL DISPARITY IN THEROPOD JAWS: COMPARING DISCRETE CHARACTERS AND GEOMETRIC MORPHOMETRICS

by JOEP SCHAEFFER , MICHAEL J. BENTON* , EMILY J. RAYFIELD  and THOMAS L. STUBBS 

School of Earth Sciences, University of Bristol, Queens Road, Bristol, BS8 1RJ, UK; mike.benton@bristol.ac.uk

*Corresponding author

Typescript received 5 February 2019; accepted in revised form 29 July 2019

Abstract: Disparity, the diversity of form and function of organisms, can be assessed from cladistic or phenetic characters, and from discrete characters or continuous characters such as landmarks, outlines, or ratios. But do these different methods of assessing disparity provide comparable results? Here we provide evidence that all metrics correlate significantly with each other and capture similar patterns of morphological variation. We compare three methods of capturing morphological disparity (discrete characters, geometric morphometric outlines and geometric morphometric landmarks) in coelurosaurian dinosaurs. We standardize our study by focusing all our metrics on the mandible, so avoiding the risk of confounding disparity methods with anatomical coverage

of the taxa. The correlation is strongest between the two geometric morphometric methods, and weaker between the morphometric methods and the discrete characters. By using phylogenetic simulations of discrete character and geometric morphometric data sets, we show that the strength of these correlations is significantly greater than expected from the evolution of random data under Brownian motion. All disparity metrics confirm that Maniraptoriformes had the highest disparity of all coelurosaurians, and omnivores and herbivores had higher disparity than carnivores.

Key words: dinosaur, coelurosaur, maniraptoriform, disparity, morphometrics.

STUDIES of the amount of morphological variation, commonly referred to as ‘disparity’, have become common in palaeontology. It might be argued that disparity (form) and diversity (species richness) should track each other in a model of homogenous evolution, but they are frequently decoupled, with some clades showing high species richness but limited diversity of form, and smaller clades showing high disparity (Wills *et al.* 1994; Fortey *et al.* 1996; Foote 1997; Ruta *et al.* 2013). Further, disparity is often high early in the evolution of a clade, suggesting some kind of ‘early burst’ model of evolution (Foote 1997; Erwin 2007). Disparity studies have provided insights into the evolution of novel body plans and ecological innovations (Goswami & Polly 2010; Brusatte *et al.* 2014; Deline *et al.* 2018), the impact and selectivity of mass extinction events (Brusatte *et al.* 2008; Friedman 2009; Bapst *et al.* 2012), and morphological expansion during evolutionary radiations (Foote 1997; Erwin 2007; Hughes *et al.* 2013; Stubbs *et al.* 2013; Close *et al.* 2015; Cooney *et al.* 2017).

Disparity should be considered in a comparative framework, and there are several analytical approaches. The

most common methods use discrete descriptive characters or geometric morphometrics. Describing morphological variation using discrete characters has usually focused on cladistic data sets as a ready source of rich data on trait variation (Wills *et al.* 1994; Lloyd 2016; Gerber 2019). This approach involves analysing character–taxon matrices where morphologies are scored using character states, including the presence and absence of features, the numbers of certain elements (e.g. teeth or limbs), the relationships between, or orientation of, elements and even general features relating to size and shape. Geometric morphometric methods, such as landmark coordinates and outlines, measure the shape of a structure, with outlines measuring the outer margin of a morphology and landmarks measuring the location of homologous features in a Cartesian coordinate system (MacLeod 1999; Zelditch *et al.* 2012).

These methods can be used in different circumstances. For some studies, geometric morphometric analyses are not possible due to a lack of homologous points, the complexity of the morphology, or a lack of completely preserved specimens, in which case discrete characters

may represent a suitable alternative. In other studies, observing shape changes, and linking these to evolutionary hypotheses, is pivotal, and discrete characters are not appropriate. Because these methods are often used interchangeably in the literature to describe disparity, it is important to consider whether all methods give similar results or not, when discussing large-scale evolutionary patterns. If not, then the methods must be applied and interpreted with extra care. Most comparisons of different disparity methods (Villier & Eble 2004; Anderson *et al.* 2011; Anderson & Friedman 2012; Foth *et al.* 2012; Hetherington *et al.* 2015; Hopkins 2017; Maclaren *et al.* 2017; Romano 2017) have shown similar results, but more or less strongly. On the other hand, Mongiardino Koch *et al.* (2017) found disagreements when comparing traditional morphometric data and discrete characters, and they strongly advocated incorporating a phylogenetic framework.

Here we compare different methods of assessing disparity, including the two main methods in geometric morphometrics (landmarks and outlines), as well as discrete cladistic characters. We use a case study focusing on a single anatomical region, to ensure that we compare like with like. Our case study looks at coelurosaurian dinosaur mandibles, and we use these for several reasons: they have a good fossil record, all clades and time intervals are sampled by multiple specimens, often with complete mandibles preserved. Further, they have a wide range of morphologies, ranging from the elongated jaws in Mesozoic birds to the robust jaws in tyrannosaurids, and the bizarre oviraptorids (Weishampel *et al.* 2004), often associated with different diets. Many coelurosaurians were carnivores, while some clades, such as therizinosaurids, oviraptorids and birds show specializations for herbivory or omnivory (Zanno & Makovicky 2011). Mandibular disparity can also be effectively measured using all analytical approaches. Finally, disparity in vertebrate jaws has been the subject of previous studies, and it is accepted that characters of the mandible and mandible shape summarize important ecomorphological traits, and variations in the morphology of the jaw are related to feeding (Anderson 2008; Anderson *et al.* 2011; Monteiro & Nogueira 2011; Stubbs *et al.* 2013; Grossnickle & Polly 2013; Zelditch *et al.* 2015; Maclaren *et al.* 2017; Hill *et al.* 2018; Nordén *et al.* 2018; Smithwick & Stubbs 2018).

MATERIAL AND METHOD

Taxon sampling

We followed two sampling approaches. In the first, a consistent sample of 40 coelurosaurian taxa was used across

all three analytical approaches, so they had identical composition in terms of phylogenetic and temporal coverage, and inter-taxon distances could be directly compared statistically. The 40 coelurosaurian taxa used in these analyses had mandibles that were complete, without taphonomic distortion and had also been coded for jaw and dental characters in the discrete character matrix of Brusatte *et al.* (2014) (see Schaeffer *et al.* 2019, tables S1, S3). Most taxa are known from a single mandible fossil, but for species where multiple specimens exist, a single representative was chosen for the geometric shape analyses, but the discrete character analyses sampled all material to maximize coding. In the second series of extended analyses, we used the maximum possible sample size available based on the restrictions of the methods. For inclusion within the discrete character analyses, the taxa had to be coded for mandibular characters in the character matrix from Brusatte *et al.* (2014), even if the specimens were fragmentary or partially incomplete. We used the function `TrimMorphDistMatrix` from `Claddis` (Lloyd 2016) to remove highly incomplete taxa that generated non-applicable distances due to a lack of shared characters, leaving a sample size of 89 taxa (see Schaeffer *et al.* 2019, tables S2, S4). For the extended geometric morphometric analyses, the jaw samples had to be complete and undistorted, but they need not have been included within the character matrix of Brusatte *et al.* (2014), giving a sample size of 60 (see Schaeffer *et al.* 2019, table S5).

Comparative groupings

We examined morphospaces and calculated disparity statistics for comparative groupings, aiming to replicate the types of analyses common in the literature. We quantified disparity in clades, firstly comparing the two major coelurosaurian groups, Maniraptoriformes and Tyrannosauroidae, and then the following subgroups within Maniraptoriformes: Dromaeosauridae, Avialae, Oviraptorosauria and Ornithomimosauria. Too few taxa were sampled from Troodontidae, Scansoriopterygidae and Therizinosauria, so these maniraptoriform subgroups were excluded from the disparity calculations. We also compared disparity in three broad ecological groupings, encompassing the dietary and size diversity within coelurosaurians: small carnivores (<4 m total body length), large carnivores (>4 m total body length) and herbivores plus omnivores. Dietary categories were assigned based on consideration of a wide range of features of anatomy and associated fossils, from the literature (Zanno & Makovicky 2011; Brusatte *et al.* 2012; Foth & Rauhut 2013). Herbivores and omnivores were grouped together because it is difficult to distinguish between

them using dietary proxies. All taxa were classified into groups using the literature (Schaeffer *et al.* 2019, tables S3–S6).

Discrete character analyses

Character-based disparity was assessed using the coelurosaurian discrete character matrix of Brusatte *et al.* (2014). First the data set was reduced to contain only mandibular and dentary tooth characters (78 characters, see Schaeffer *et al.* 2019, tables S1, S2). As previously noted, we first ran analyses with a reduced sample to match the geometric morphometric analyses (40 taxa with complete mandibles), and then additional analyses were run using the full sample. The inclusion of dentary teeth characters within the analyses was assessed by running supplementary analyses using a mandible-only data set (66 characters).

The character matrices were converted into pairwise distance matrices based on maximum observable rescaled distances (MORD; Lloyd 2016). The distance matrices were then subjected to principal coordinates analysis (PCOA) with Cailliez negative eigenvalue correction (Cailliez 1983), and the latter had the effect of reducing the proportion of variance expressed by each PCO-axis compared to uncorrected PCOA (i.e. PCO1 10.7% vs 25.1% and PCO2 9.6% vs 21.3%) (Schaeffer *et al.* 2019, figs S1, S2). However, the distribution of taxa in morphospace is almost identical in biplots generated from the major axes of variation (PCO1–PCO5), as previously noted by Hopkins (2017) and Nordén *et al.* (2018). Therefore, although the percentage of variance reported is comparatively low, the biplots of PCO1–PCO2 morphospace still represent a decent proportion of overall variance and illustrate the major axes of variation. The disparity metric sum of variances (SOV) was calculated from all coordinate axes scores for each of the comparative groupings. The SOV metric is generally robust to sample size differences in comparative groupings (Ciamaglio *et al.* 2001), but to test this we also calculated SOV with rarefaction to equalize sample sizes based on the group with smallest sample size. All analyses were performed using the packages *phytools* (Revell 2012), *Claddis* (Lloyd 2016), *disprity* (Guillerme 2018) and *vegan* (Oksanen *et al.* 2017), all in the R coding environment (R Core Team 2017).

Geometric morphometric analyses

The geometric shape analyses were based on lateral profiles of the mandibles in two dimensions. We first assembled a collection of lateral mandible images from museum specimens and the primary literature (Schaeffer

et al. 2019, table S8). For the outline analyses, the jaw samples were converted to silhouettes, excluding the dentition, and imported into R. We used elliptical Fourier analysis (EFA) to compare outline shapes with the R package *Momocs* (Bonhomme *et al.* 2014; Navarro *et al.* 2018). The jaws were converted to digitized outline closed curves (Fig. 1C), and then aligned, centred and scaled, before a set of 500 *x* and *y*-coordinate points was assigned to each outline profile. The number of harmonics required to account for the shape variation was calibrated. Each harmonic yielded four Fourier coefficients describing the shapes. Fourier coefficient data accounting for 99% of the shape variation in the original jaws (here coefficients from 11 EFA harmonics) was subjected to principal component analyses (PCA) to ordinate the data, explore major aspects of the geometric variation and generate morphospaces using the first two principal components axes. The same jaw samples were used in the landmark geometric analyses. Landmarks were digitally added to the specimens using *tpsDig2* and *tpsUtil* (Rohlf 2017a, b) and consisted of 6 fixed landmarks and 50 semi-landmark points along 6 curves (Bookstein 1991) (Fig. 1B, Schaeffer *et al.* 2019, files S1–S4). To remove noise effects such as size and orientation, Procrustes aligned landmarks were calculated using a generalized Procrustes analysis in *tpsRelw*, that incorporated a sliding procedure for minimizing bending energy (Rohlf 2017c). The Procrustes-aligned landmark data were then subjected to PCA in the R package *geomorph* (Adams & Otárola-Castillo 2013), and morphospaces were created using the scores from the first two axes. For both geometric morphometric analyses, the function *PCcontrib* from the R package *Momocs* (Bonhomme *et al.* 2014) was used to visualize and plot shape changes along principal component axes 1, 2 and 3, and the sum of variances disparity statistic was calculated for the comparative groupings using all morphospace axes scores with the R package *disprity*, both with and without rarefaction (Guillerme 2018).

Correlation tests

Morphological disparity can be expressed as pairwise distances between taxa. To assess whether the three measures of disparity, based on characters, landmarks and outlines, provide similar insights into morphological disparity, we used Mantel tests (Anderson & Friedman 2012; Hetherington *et al.* 2015) to examine the correlation between pairwise distances derived from each sample-standardized analysis. Pairwise distances for the three data sets are based on Euclidean distances in the multidimensional morphospace coordinates from all axes, generated from the protocols described above. Ancillary tests were

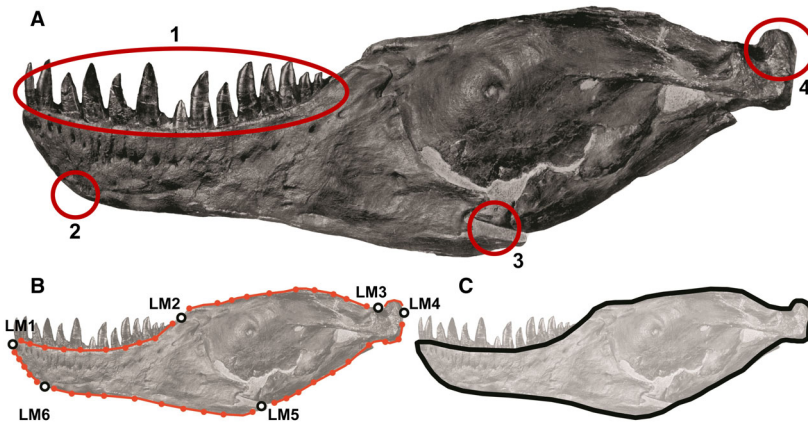


FIG. 1. Three approaches for measuring morphological disparity. The mandible of *Tyrannosaurus rex* (AMNH 5027) is used to illustrate the three analytical approaches. A, examples of discrete jaw and dental characters, namely: (1) the presence or absence of dentary teeth; (2) the shape of the anteroventral angle of the dentary; (3) the external mandibular fenestra location; and (4) the retroarticular process, presence and shape. B, six landmarks and 50 semi-landmarks digitized on a *T. rex* mandible; the fixed landmarks are: LM1, the anterior-most point of the mandible at its dorsal edge (Type 2), taken as the dorsal most tip of the anterior portion of the jaw; LM2, the dentary–surangular suture at the dorsal edge of the mandible (Type 1); LM3, in the articular bone, the centre of the glenoid (Type 3); LM4, the most posterior point of the mandible (Type 2), taken from the angular or the surangular bone, depending on which element was the most posterior; LM5, the dentary–angular suture at the ventral edge of the mandible (Type 1); LM6, the anterior-most point of the mandible at its ventral edge (Type 2), taken as the ventral most tip of the anterior portion of the jaw where the angle of jaw symphysis changes. C, geometric morphometric outline illustrated on the *T. rex* mandible. Note that the dentition is not included in the landmarks or outlines. Colour online.

performed based on pairwise distances from character data with dentary teeth characters excluded, to ensure that the inclusion or exclusion of teeth in the character data was not significantly impacting the measured disparity.

Phylogenetic simulation tests

To provide a null model with which to compare our empirical correlation results we used a simulation approach. We tested whether the empirical correlation results between different data sets describing the same anatomical unit are significantly different to correlations expected from the evolution of unrelated traits or morphological data sets evolving under Brownian motion on a phylogeny.

For the phylogenetic data we sampled 10 coelurosaurian trees from Brusatte *et al.* (2014). To calculate phylogenetic branch lengths, each topology was time-calibrated 50 times using both the equal (Brusatte *et al.* 2008) and minimum branch length (Laurin 2004) dating approaches with the `timePaleoPhy` function of the `paleotree` R package (Bapst 2012). For each dating replicate, ages were randomly sampled from between each taxon's first and last appearance dates (age data from Brusatte *et al.* 2014). The resulting 1000 coelurosaurian trees therefore incorporated phylogenetic uncertainty (10 topologies),

two dating methods, and temporal occurrence uncertainty for the tips (through the replicates). All trees were cropped to contain only the 40 taxa that were present in the empirical correlation tests. We then simulated two types of morphological data, with similar properties to those in our study, for each of the 1000 dated phylogenies.

First, we simulated discrete character data sets. The function `sim.morpho` from the R package `disparity` (Guillaume 2018) was used to simulate data sets with 78 characters (77% of which had two states and 23% had three states) using the equal-rates (ER = Mk) model (Paradis & Schliep 2018). For the model parameters of each simulated data set, we randomly sampled rates of 5, 10 or 20 (higher rates increased phylogenetic signal) and distribution shapes of 0.5, 1 or 2. Missing data, ranging from 21.5% to 37.8%, was introduced to each simulated discrete character data set to reflect the nature of the empirical jaw character data (and fossil-based discrete character data in general). The simulated discrete character data were converted into MORD pairwise distance matrices and subjected to PCOA with Cailliez negative eigenvalue correction (Cailliez 1983), giving 1000 sets of PCOA scores for 40 taxa.

Next, we simulated geometric shape data evolving under Brownian motion for the same 1000 dated coelurosaurian trees. We aimed to generate 40 'jaw-like' structures defined by 56 landmark coordinates for each tree. As a starting

point, we used the consensus shape from our sample of 40 theropod jaws as the ancestral morphology. The method required co-variance data to guide how the shapes would transform during the simulations, so we calculated Procrustes residuals from our original landmark configurations as input. With these three inputs, a phylogeny, ancestral morphology and residuals data, we simulated landmark configurations at all nodes and tips evolving under phylogenetic Brownian motion for all 1000 topologies, using custom code built around the SimEvo function from Evomorph (Cabrera & Giri 2016). This model simulates shapes by calculating the product of a phenotypic covariance matrix (which is the Procrustes residuals not the phylogenetic variance co-variance matrix) with the vector βH . The H component transforms the P matrix into the G matrix which describes the heritable proportion of P (Polly 2004), with β summarizing the selection coefficients (i.e. morphological change over time = $\beta * H * P$). This is identical to the formulation in SimEvo, except that instead of running the simulation for n generations, our method runs a single generation in which variation of $\beta * H$ is sampled according to a mean zero normal distribution with variance equal to its branch length. As a Brownian process evolution occurs independently on each branch, with each descendant branch inheriting the shape from its parent. After running the simulations, the 1000 sets of 40 landmark configurations were converted to TPS format and subjected to PCA in *geomorph* (Adams & Otárola-Castillo 2013), giving 1000 sets of PC scores (code is available in Schaeffer *et al.* 2019).

Mantel correlation tests were performed for each pair of simulated data sets (characters and landmarks simulated from the same tree) based on pairwise Euclidean distances from the multidimensional morphospace coordinates from all axes. The distribution of correlation statistics from these 1000 paired simulated data sets was then statistically compared to the observed correlation results from the jaw discrete character and geometric shape data sets, to test the null hypothesis of no difference between the simulated data sets and real data.

RESULTS

Morphospace occupation

The morphospace biplots (Figs 2, 3) illustrate the morphological dissimilarity between taxa based on major features of morphological variation and they show divisions according to taxonomic grouping (Fig. 2) and dietary category (Fig. 3). In both figures, the morphospace axes represent the same major features of variation, so this is described first, before exploring how the taxonomic and dietary groupings are distributed in morphospace.

In the outline morphospaces, using the standardized sample of 40 jaws (Figs 2A, 3A), PC1 describes 61.4% of variation, representing the thickness of the dentary region and jaw mid-length, while PC2 (23.2%) encapsulates variation in the dorsoventral curvature of the mandible and the relative position of the cranium–mandible articulation (Fig. 4). The remaining outline morphospace axes (PC3–PC40) describe 15.4% of variation (Schaeffer *et al.* 2019, fig. S3). In the standardized landmark analyses (Figs 2B, 3B, 4), PC1 (40.7%) reflects changes in jaw depth, particularly at the mid-length and anteriorly, like the outline analysis. However, PC2 (22.8%) represents the overall thickness of the jaw, unlike the outline analysis. Variation in the dorsoventral curvature of the jaw is represented by PC3 (13.6%) in the landmark analysis (Schaeffer *et al.* 2019, fig. S4). The remaining axes in the landmark analysis (PC4–PC40) describe a further 22.9% of variation (Schaeffer *et al.* 2019, fig. S4). In the standardized discrete characters analyses (Figs 2C, 3C), PCO1 accounts for 10.7% of variation while PCO2 represents a further 9.6% (Schaeffer *et al.* 2019, fig. S2). As previously noted, these relatively low percentages of variation are the combined result of having many characters, missing data and using negative eigenvalue correction during PCOA (Hopkins 2017; Nordén *et al.* 2018).

In the morphospaces highlighting taxonomic groupings we compare the distribution of maniraptoriforms and tyrannosauroids (Fig. 2). In all analyses, the former occupies a much larger region of morphospace than the latter. The jaw outlines morphospace (Fig. 2A) shows a near-complete separation between maniraptoriforms and tyrannosauroids. Three basal tyrannosauroids, *Proceratosaurus*, *Guanlong* and *Dilong*, overlap with maniraptoriforms in central morphospace. The maniraptoriforms occupy a large range along PC1, mainly due to the divergent oviraptorids, and a smaller range along PC2, whereas the tyrannosauroids have a large range on PC2 but limited expanse on PC1. The landmarks morphospace (Fig. 2B) shows complete overlap of the two groups. The maniraptoriforms occupy a broad area that is extended along both PC1 and PC2, whereas tyrannosauroids have limited expanse on both PC1 and PC2. The discrepancy between the two morphometric morphospaces results from different aspects of shape variation being represented by PC2 (Fig. 4). If landmark morphospace is plotted based on PC1 and PC3, the overall pattern is similar to the outline morphospace, with tyrannosauroids diverging on PC3 (Schaeffer *et al.* 2019, fig. S7). The discrete characters morphospace (Fig. 2C) shows similar patterns to the outlines morphospace (Fig. 2A), with two basal tyrannosauroids, *Proceratosaurus* and *Guanlong*, overlapping with maniraptoriforms, while tyrannosauroids generally occupy a distinct area of morphospace and a restricted range of PCO1, but expand on PCO2. The

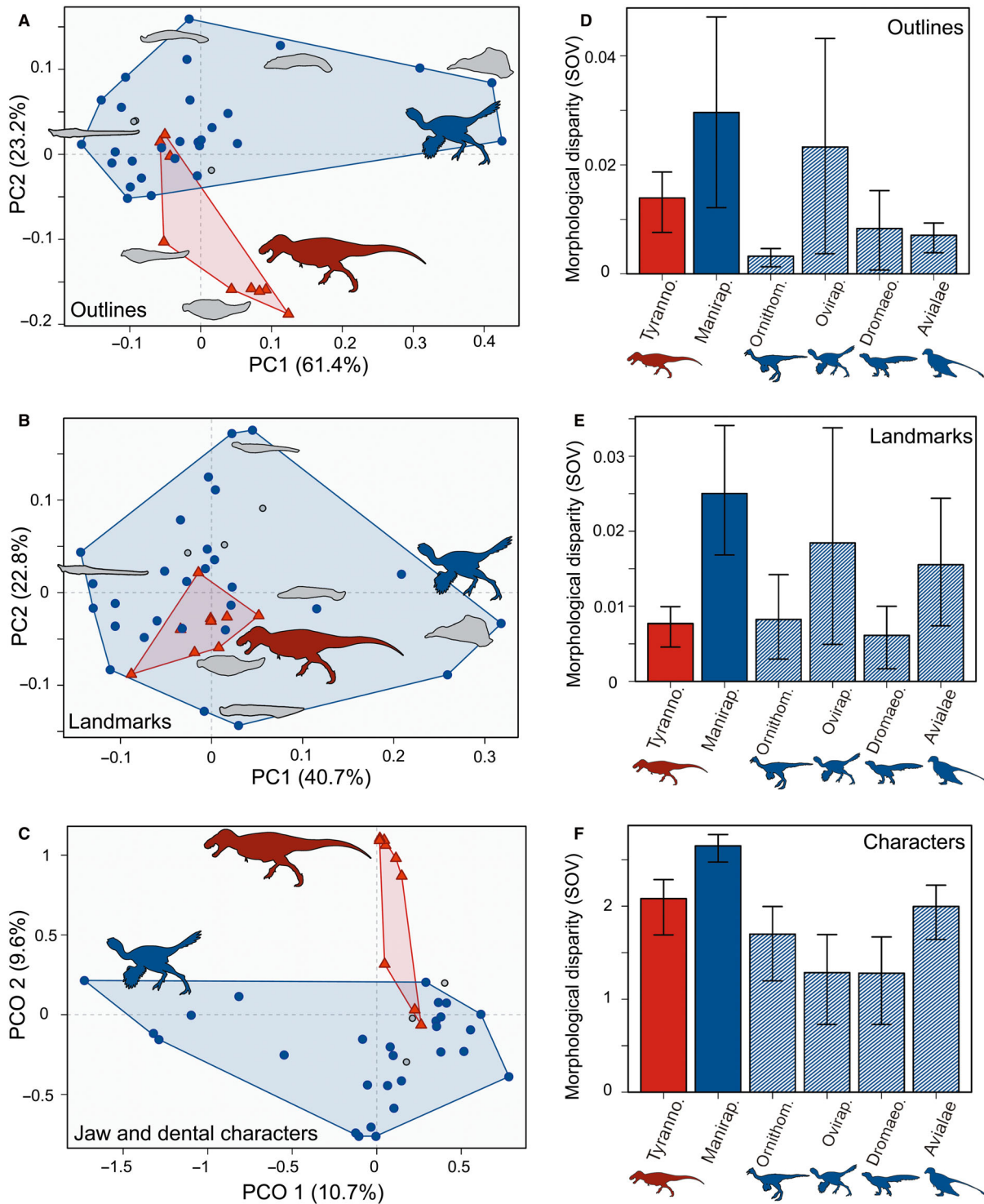


FIG. 2. Patterns of morphospace occupation and disparity (sum of variances, SOV) of Coelurosauria grouped according to taxonomy, measured using three metrics. The bivariate morphospaces and disparity plots were created using: A, D, geometric outlines; B, E, geometric landmarks; C, F, discrete jaw and dental characters. *Legend:* triangles, Tyrannosauroidae; circles, Maniraptoriformes; stripes (D–F), disparity for the Ornithomimosauria, Oviraptoridae, Dromaeosauridae and Avialae. In the morphospace plots (A–C) open circles denote coelurosaurians that are not classified as Tyrannosauroidae or Maniraptoriformes, namely *Sinosauropteryx prima*, *Compsognathus longipes* and *Ornitholestes hermanni*. Dinosaur silhouettes from <http://phylopic.org>: Scott Hartman, Nobu Tamura, Brad McFeeters, Emily Willoughby and Matt Martyniuk (CC BY 3.0). Colour online.

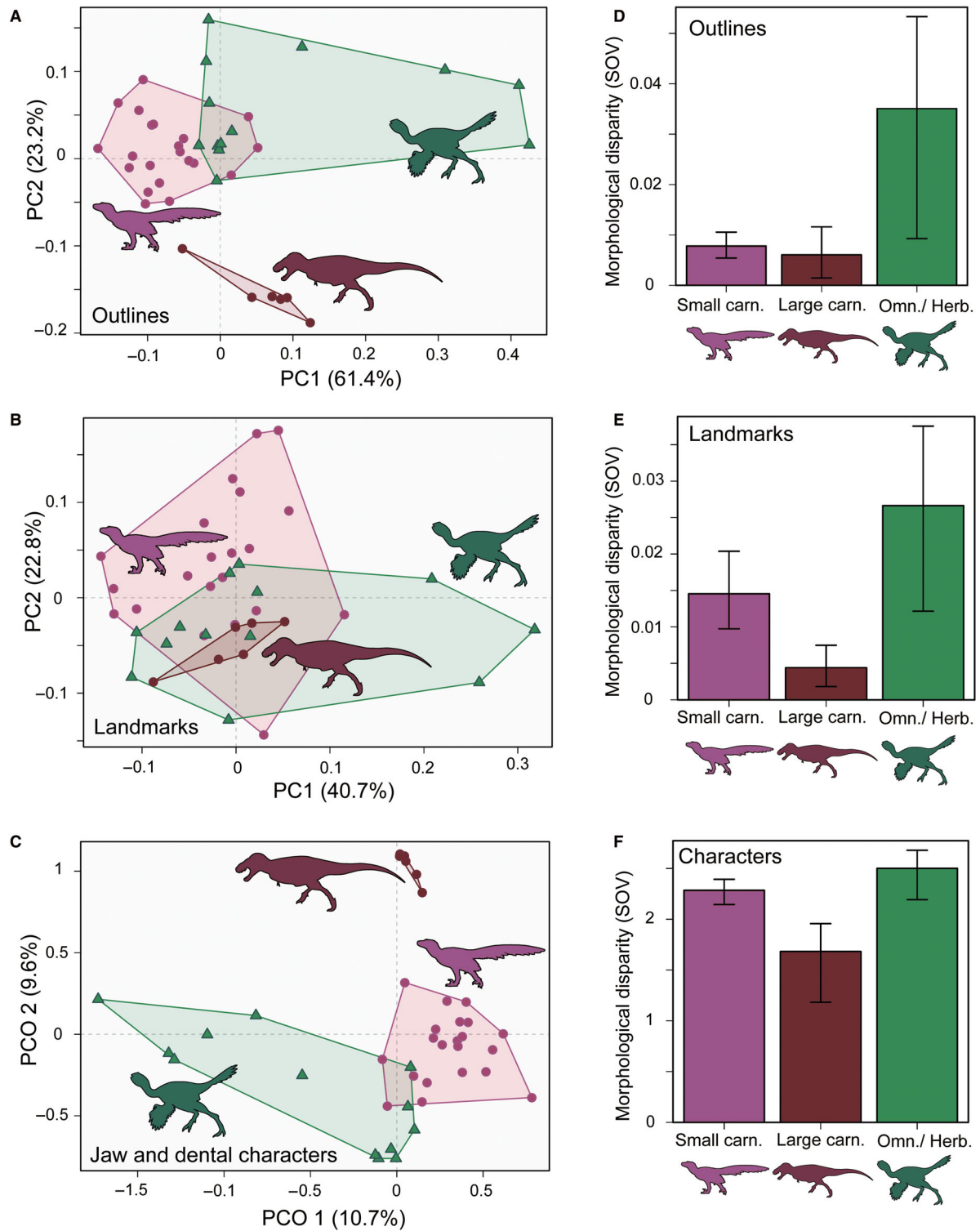


FIG. 3. Patterns of morphospace occupation and disparity (sum of variances, SOV) of Coelurosauria grouped according to diet, measured using three metrics. The bivariate morphospaces and disparity plots were created using: A, D, geometric outlines; B, E, geometric landmarks; C, F, discrete jaw and dental characters. *Legend:* circles, small carnivores; dark circles, large carnivores; triangles, omnivores and herbivores. Dinosaur silhouettes from <http://phylopic.org>: Scott Hartman, Brad McFeeters and Emily Willoughby (CC BY 3.0). Colour online.

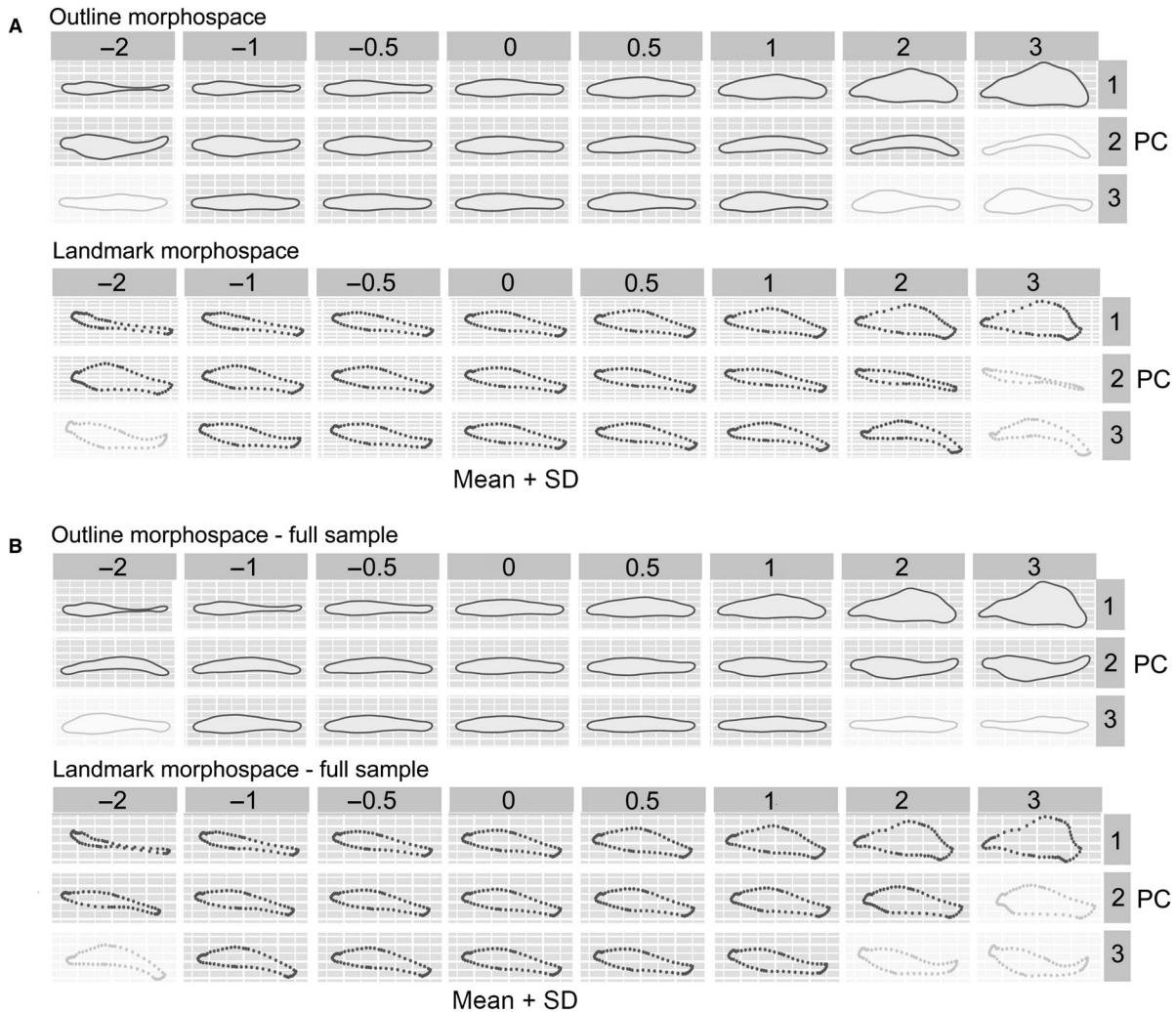


FIG. 4. Mandible shape changes along the three main principal components axes (PC1, PC2, PC3). A, shape changes based on outline and landmark approaches for the standardized sample ($n = 40$). B, shape changes based on outline and landmark approaches for the full sample ($n = 60$). The y -axes are the principal component axes. The x -axes illustrate the mean shape and the shapes at standard deviations for each of the PC-axes. Darker outlines are realized morphologies whereas lighter outlines are theoretical morphologies not shown by the sampled taxa.

maniraptoriforms occupy a wide zone, particularly extending along PCO1 due to the divergent oviraptorids. Morphospace plots with all taxa individually labelled are provided in Schaeffer *et al.* (2019, figs S5, S6, S8, S9). When using a reduced discrete character data set of 66 characters, based only on the mandibular bones (excluding teeth), the overall distribution of taxa is similar, as is the relative expanse of both clades on the major PCO axes (Schaeffer *et al.* 2019, fig. S10A).

In the morphospace plots showing the taxa distinguished by diet there are many commonalities in the distribution of groups and some nuanced differences (Fig. 3). No single morphotype or discrete character combination characterizes omnivores and herbivores, but instead the grouping is

widely distributed in all morphospaces, occupying extreme positions on the major axes of variation and more central areas. The wide distribution in morphospace probably reflects the fact that this is an ill-defined dietary category. Some of the sampled small carnivores overlap with omnivore–herbivores in central morphospace, but the grouping does contain some divergent forms, particularly in the landmark morphospace on PC2 (Fig. 3B). The sampled large carnivorous taxa occupy a reduced area of morphospace in all analyses. Their morphospace is distinct from the small carnivores and omnivore–herbivores in the outline and character morphospaces, but not in the landmark morphospace (Fig. 3B). Large carnivorous taxa generally have jaws that are dorsoventrally deep in the

postdentary region. Once again, when using a reduced discrete character data set with only mandibular bones the relative distribution and expanse of dietary groups in morphospace is consistent (Schaeffer *et al.* 2019, fig. S10B).

Disparity patterns

As with the morphospaces, the sum of variances (SOV) disparity patterns converge for the different metrics but show some subtle differences between the data types (Figs 2, 3). In the taxonomic results (Fig. 2), the tyrannosauroids always have lower mandibular disparity than the maniraptoriforms. These results are statistically significant, based on non-overlap of the 95% confidence intervals generated through bootstrapping, for the landmarks and characters (Fig. 2E, F, Schaeffer *et al.* 2019, fig. S10C) but not the outlines (Fig. 2D). The confidence intervals around the outline-based SOV statistic for maniraptoriforms are large, probably owing to the inclusion/exclusion of certain divergent oviraptorid taxa during the bootstrapping procedure. Within Maniraptoriformes, the greatest disparity in the outline analysis (Fig. 2D) is seen in oviraptorids, and a smaller portion from the Dromaeosauridae, Avialae and Ornithomimosauria. The landmarks results (Fig. 2E) show a similar pattern, with the oviraptorids showing greatest disparity, but this time closely followed by the Avialae.

The character-based results (Fig. 2F, Schaeffer *et al.* 2019, fig. S10C) show something completely different. Within the Maniraptoriformes, the Avialae show the highest subgroup disparity, followed by the ornithomimosaurids, oviraptorids and dromaeosaurids, but all are within the error bar ranges of each other. The dietary groupings also show some differences between the three methods (Fig. 3). In all three cases, the omnivore–herbivore grouping shows highest disparity, but this is notably higher in the shape analyses (Fig. 3D, E) and only slightly so according to discrete characters (Fig. 3F) reflecting the fact that the category may include several feeding types that we cannot further subdivide. Further, the large carnivores always show lowest total disparity, but significantly so only for landmarks (Fig. 3E) and discrete characters (Fig. 3F, Schaeffer *et al.* 2019, fig. S10D). All disparity results are consistent when using rarefaction to standardize sample sizes, but the confidence envelopes are wider due to reduced sample sizes (Schaeffer *et al.* 2019, table S9).

Correlation between morphological distances

Across all correlation tests based on morphological distances, both Pearson and Spearman Mantel tests show

evidence of statistically significant correlation at a 0.001 threshold level. The strength of correlations varies between tests. As expected, from the two geometric morphometric approaches, the outline and landmark-based distances, have a strong and significant correlation (Pearson $r = 0.791$, $r^2 = 0.626$, $p = 0.001$; Spearman $\rho = 0.579$, $p = 0.001$). The strength of correlations between morphological distances in the discrete character morphospace and the shape morphospaces are weaker. The discrete character data does show a relatively strong correlation with the outline distances (Pearson $r = 0.611$, $r^2 = 0.373$, $p = 0.001$; Spearman $\rho = 0.7147$, $p = 0.001$), while the weakest correlation between distances is recovered from tests of the discrete character distances and the landmark distances (Pearson $r = 0.502$, $r^2 = 0.252$, $p = 0.001$; Spearman $\rho = 0.491$, $p = 0.001$). When dental characters are excluded from the discrete character analyses, the correlation tests again show statistically significant correlations with the outline data (Pearson $r = 0.578$, $r^2 = 0.334$, $p = 0.001$; Spearman $\rho = 0.7278$, $p = 0.001$) and weaker correlation with the landmark data (Pearson $r = 0.452$, $r^2 = 0.205$, $p = 0.001$; Spearman $\rho = 0.4591$, $p = 0.001$). The correlation between morphological distances in discrete character morphospaces with and without dental characters is very strong and significant (Pearson $r = 0.964$, $r^2 = 0.929$, $p = 0.001$; Spearman $\rho = 0.970$, $p = 0.001$), suggesting the addition of 13 dentary tooth characters does not alter the character-based disparity patterns.

Correlations and phylogenetic simulations

Correlation tests performed on the phylogenetically simulated discrete character and geometric shape data also return statistically significant results, but the correlation coefficients and coefficients of determination are significantly lower than in the empirical jaw data sets. Mantel tests performed on the 1000 pairs of phylogenetically simulated discrete character data and shape data show that 81.1% (Pearson) and 90% (Spearman) of iterations give statistically significant correlations at a 0.05 threshold level, while 67.3% (Pearson) and 79.7% (Spearman) return statistically significant correlations at a 0.001 threshold. However, when comparing the distribution of correlation coefficients and coefficients of determination from the simulated results to the empirical results, it is clear that the Pearson's r values (outlines vs characters $p < 0.001$; landmarks vs characters $p = 0.001$), r^2 values and Spearman's ρ values (outlines vs characters $p < 0.001$; landmarks vs characters $p = 0.016$) are significantly higher in the real mandibular data than in the correlations from 1000 phylogenetically simulated character and shape data sets (Fig. 5).

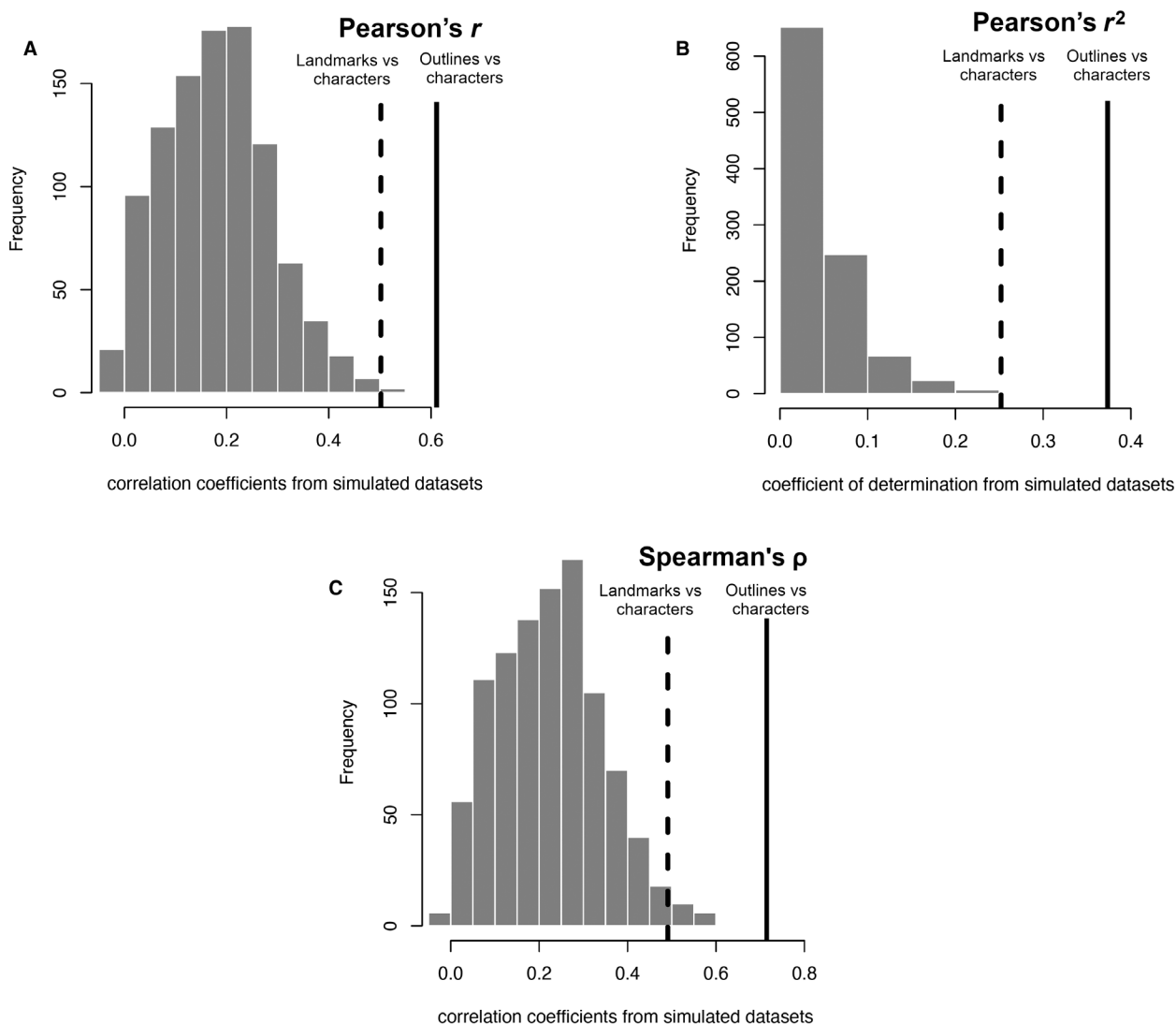


FIG. 5. Correlation results from phylogenetically simulated discrete character and shape data compared to the empirical correlation coefficients (A, C) and coefficients of determination (B). Histograms show the distribution of results from the simulated data correlations. Dashed lines denote the correlation coefficients and coefficients of determination for jaw landmarks vs discrete characters, and the solid lines show the results from correlation tests between jaw outlines and discrete characters.

Full sample size

For the full sample (Figs 6, 7) the same disparity patterns are present as in the standardized sample (Figs 2, 3). The relative distributions of tyrannosauroids and maniraptoriforms remain similar for all analyses (Fig. 6A–C), and the relative proportions of morphospace occupied by all clades remain comparable. In the two shape analyses, the same geometric changes are recovered on PC1–PC3 in the full sample size analyses (Fig. 4; although note that PC2 and PC3 are reversed), suggesting the inclusion of additional samples does not alter our understanding of the major shape innovations in coelurosaurian jaws. Expanding the sample for the geometric analyses introduced two outlying

forms; the bird *Longipteryx* notably diverges along PC2 in the outline morphospace and the oviraptorid *Gigantoraptor* expands the bounds of PC2 in the landmark morphospace. When using the full sample, the maniraptoriforms again show significantly greater SOV disparity than the tyrannosauroids, but now in all three analyses (Fig. 6D–F), while the Avialae show relatively higher disparity in the outline analysis when compared to the standardized sample (Fig. 6D). In the discrete character analysis, the various maniraptoriform subclades show much more uniform total disparities for the full data set (Fig. 6F) than for the standardized data set (Fig. 2F).

In terms of dietary categories, the morphospace and disparity patterns are also similar (Figs 3, 7). Omnivores–

herbivores again show the greatest disparity in all analyses. The disparity of large carnivores is greater when using the full sample, and the group is represented by 11 taxa rather than 6 taxa. This is particularly noticeable in the extended discrete characters study (Fig. 7C, F), where large carnivores now occupy a larger area of morphospace and overlap the small carnivores when more taxa are included. The relative proportions of the total morphospaces are similar between the reduced taxon study (Fig. 3D–F) and the study with all taxa (Fig. 7D–F), except for the relative levelling of total SOV for all three categories when sampling is increased (cf. Figs 3F, 7F). All disparity results are consistent when using rarefaction to standardize sample sizes (Schaeffer *et al.* 2019, table S9).

DISCUSSION

Morphological disparity and a comparison of methods

The three methods of measuring morphological disparity share many commonalities and show moderate correlations between morphological distances. By incorporating a simulation approach, we tested whether the observed correlation results should be expected from random evolution under phylogenetic Brownian motion, and we used a novel modelling approach to simulate geometric data. It is widely accepted that shared phylogenetic history leads to phenotypic similarity between related species (Felsenstein 1985; Harvey & Pagel 1991; Freckleton *et al.* 2002). It is therefore unsurprising that our tests recover significant correlations between the majority of phylogenetically simulated shape and discrete character data sets, highlighting a common phylogenetic structure. However, the strength of the correlations between the three coelurosaur jaw data sets could not be replicated by simulations and the contrasting measures of morphological disparity are significantly more similar than expected from the evolution of random data under Brownian motion. Given that the shape analyses and discrete characters record morphological differences in very different ways, even if measured from the same anatomy, the moderate to strong correlations suggest that the three methods capture the same major patterns of morphological disparity, and this is not solely the outcome of random evolution along the branches of a phylogeny.

In light of our results, it is important to consider the extent to which we should expect the geometric shape and discrete character data to capture the same disparity patterns. Of the 78 jaw and dental characters used, 66 specifically relate to the mandible. However, only 28.8% of these characters (19 characters: numbers 1, 2, 3, 6, 8, 9, 14, 26, 28, 37, 42, 43, 44, 45, 47, 51, 61, 72, 78)

describe overall shape variation, and just 18.2% of the 66 characters (12 characters: numbers 2, 3, 6, 14, 26, 28, 43, 45, 47, 51, 61, 78) could be captured by 2-D lateral jaw profiles in our analyses (other variation is in the dorsoventral shape or fenestrae shape/position). The remaining discrete mandibular characters encompass more nuanced morphological features, such as sutural contacts between constituent bones, presence, size and location of foramina and fossae, bone textures, pneumaticity, and features of muscle attachment sites. Therefore, it is perhaps surprising that we recover moderate correlations between the shape and character data. This may result from concerted evolution between geometry and other characters of the mandible, potentially highlighting morphological integration, an ecological constraint, or an underlying phylogenetic signal (Hetherington *et al.* 2015). Furthermore, it is surprising that correlations between the geometric data and discrete character data are stronger when teeth are included (all 78 characters), compared to comparisons with only the mandibular bones (66 characters). This implies that taxa with divergent jaw shapes share dental character scoring, and therefore the inclusion of dental characters does not erode the major dissimilarity patterns. Again, this could represent an ecological characteristic or an underlying phylogenetic signal.

Our study can be compared with previous endeavours, which have so far yielded differing results. Hetherington *et al.* (2015), for example, working on caecilian amphibians, showed significant and moderate to good correlations (Spearman's $\rho = 0.36$ – 0.66 and Pearson $r = 0.38$ – 0.65 , for comparisons of Euclidean distance matrices from the different methods) between landmark-based morphometrics and discrete character methods, confirming earlier results on echinoderms by Villier & Eble (2004). On the other hand, Mongiardino Koch *et al.* (2017), in their study of scorpion disparity, found significant differences between discrete character and traditional morphometric methods, and they showed that these differences were greater than expected from the evolution of random data on a phylogeny. They also simulated discrete character data and compared this to simulated traditional morphometric data (not landmark coordinates).

As expected, the two shape-based methods (outlines and landmarks) show the greatest similarity, both in terms of the relative areas of morphospace occupation, the overlaps in morphospace, the total disparity measured by SOV, and in correlation tests. The similarities between the outline and landmark measures are readily understood when shape transformations are compared using both methods and both sample sizes for the first three principal components (Fig. 4). In both cases, the PCA is capturing similar aspects of variation: jaw height, relative attenuation along the jaw length and jaw curvature. The

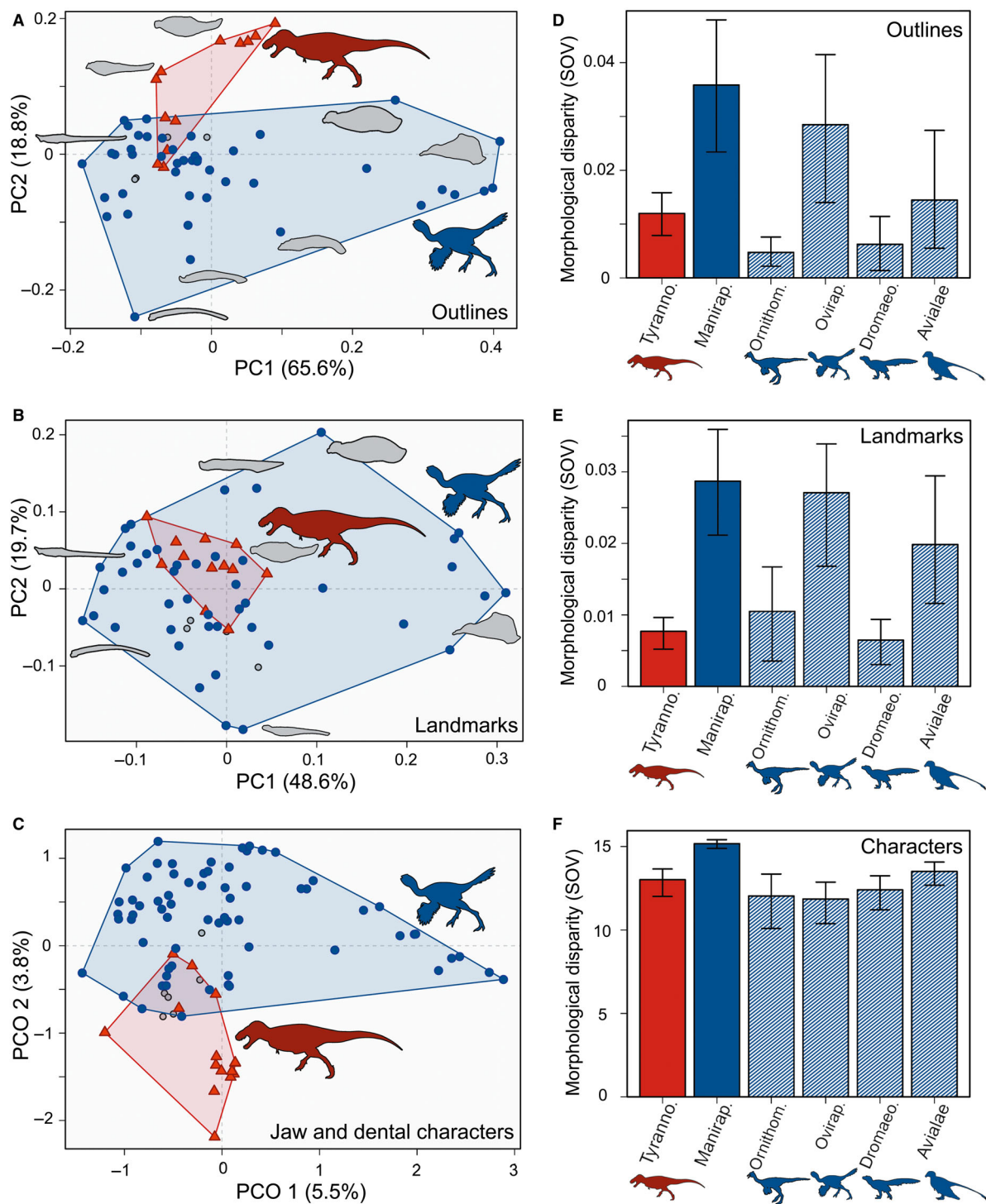


FIG. 6. Patterns of morphospace occupation and disparity (sum of variances, SOV) of Coelurosauria, for the full sample, grouped according to taxonomy, measured using three metrics. The bivariate morphospaces and disparity plots were created using: A, D, geometric outlines; B, E, geometric landmarks; C, F, discrete jaw and dental characters. *Legend:* triangles, Tyrannosauroidae; circles, Maniraptoriformes; stripes (D–F), disparity for the Ornithomimosauria, Oviraptoridae, Dromaeosauridae and Avialae. In the morphospace plots (A–C) open circles denote coelurosaurians that are not classified as Tyrannosauroidae or Maniraptoriformes, these taxa are listed in Schaeffer *et al.* (2019, tables S4, S5). Dinosaur silhouettes from <http://phylopic.org>: Scott Hartman, Nobu Tamura, Brad McFeeters, Emily Willoughby and Matt Martyniuk (CC BY 3.0). Colour online.

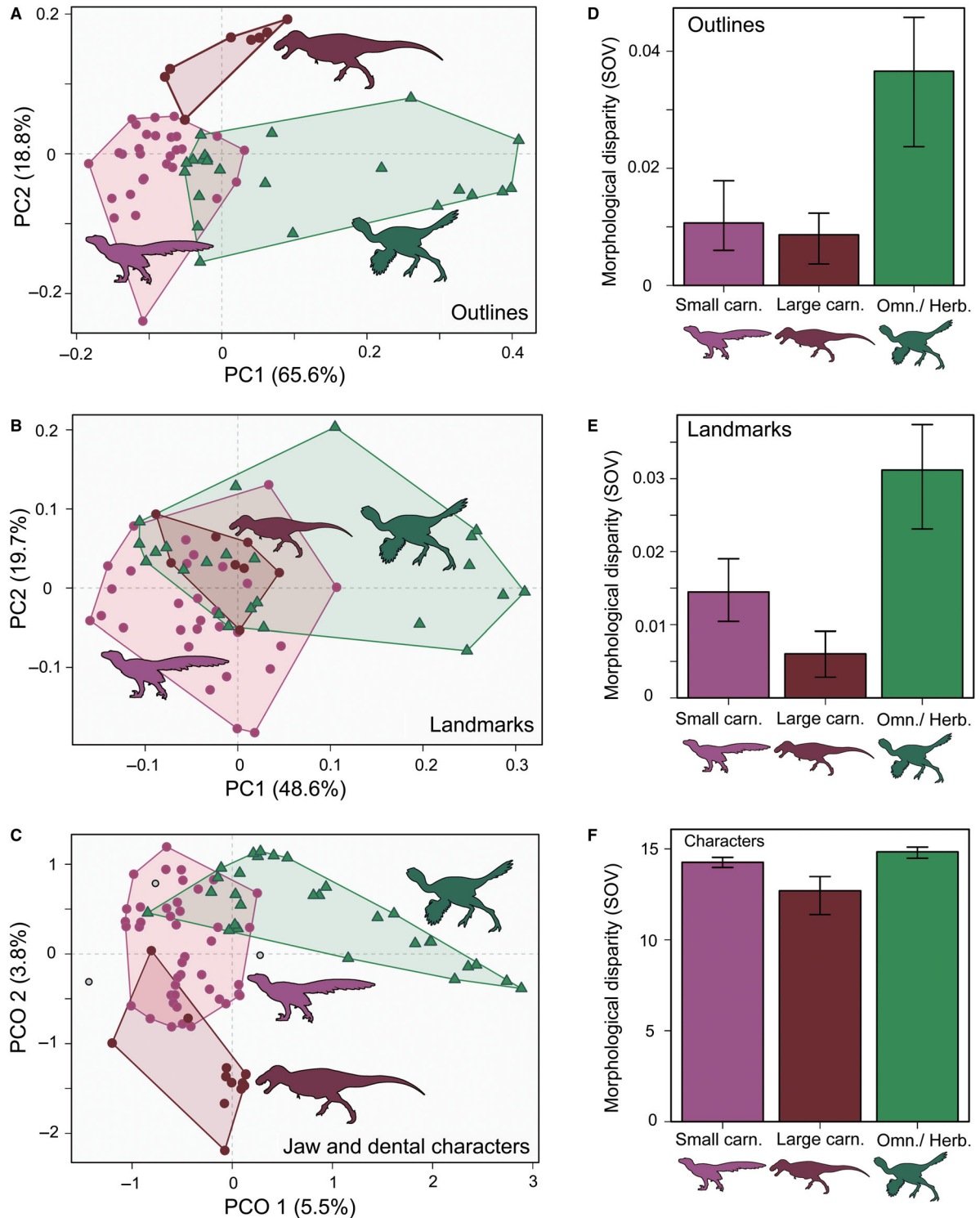


FIG. 7. Patterns of morphospace occupation and disparity (sum of variances, SOV) of Coelurosauria, for the full sample, grouped according to diet, measured using three metrics. The bivariate morphospaces and disparity plots were created using: A, D, geometric outlines; B, E, geometric landmarks; C, F, discrete jaw and dental characters. *Legend:* circles, small carnivores; dark circles, large carnivores; triangles, omnivores and herbivores; open circles (C), coelurosaurians that could not be confidently assigned a dietary grouping, these taxa are listed in Schaeffer *et al.* (2019, tables S4, S5). Dinosaur silhouettes from <http://phylopic.org>: Scott Hartman, Brad McFeeters and Emily Willoughby (CC BY 3.0). Colour online.

differences between the two shape analyses are because PC2 and PC3 variously capture the concavity of the jaw outline.

The discrepancy between PC2 in the shape analyses may be driven by our sampling of landmarks. In the landmark geometric approach, two fixed landmarks were positioned on the anterior part of jaw, at the dorsal and ventral angles of the symphysis, and a landmark curve with five points was located between these fixed points. Morphological changes in the relative locations of these anterior landmarks were dominant factors loading on PC2 in the landmark analyses and resulted in the dorsoventral curvature being expressed on PC3, instead of PC2 as in the outline analyses (Fig. 4). It could therefore be argued that this region of the jaw was oversampled, while the outlines more effectively sampled the jaws edges with densely sampled, but evenly distributed, points. Landmarking procedures are an important part of shape analyses (Mitteroecker & Gunz 2009; Cardini 2016; Watanabe 2018) and our results highlight how the selection of methods and landmarks could impact the interpretations of shape evolution based on the major axes of variation. Although, it is important to note that the SOV disparity measured from all PC axes (total shape variation) gave consistent patterns in the outline and landmark methods and the correlation tests based on all axes show strong and significant correlations.

Perhaps our application of landmarks tended to increase the similarity of the outlines and landmark results. By adopting multiple richly sampled semi-landmark curves, we effectively outlined each mandible with 50 points, which gives nearly as much detail as the outline approach (Figs 1, 4). Nonetheless, we felt it was inappropriate here to use a reduced number of fixed homologous landmarks, because there were only a limited number of confidently identifiable homologous points and introducing semi-landmark curves more accurately captures additional information on the overall shape of the mandible. Ideally, it would have been interesting to incorporate landmarks measuring the size and location of the functionally significant external mandibular fenestra, but this feature is absent in some Mesozoic birds and therefore could not be measured on all samples.

The SOV results all suggest that Maniraptoriformes are more disparate than Tyrannosauroida (Figs 2, 6). The relatively low disparity of Tyrannosauroida could reflect the lower diversity of this clade in comparison to the Maniraptoriformes, but we argue this is not a problem of experimental design. First, we performed additional analyses with rarefaction to standardize sample sizes for SOV disparity comparisons, and the results were consistent with unrarefied analyses (Schaeffer *et al.* 2019, table S9). Second, we included all possible specimens (Fig. 6), and so the differences reflect reality, namely the fact it is a

smaller clade and with more uniform anatomy and adaptation than Maniraptoriformes. Finally, we used Brusatte *et al.* (2014) as the data source for discrete cladistic characters, and those authors included the maximum possible sample of tyrannosauroid taxa compared to the other groups (such as Avialae); this does have the effect of diminishing the difference between Maniraptoriformes and Tyrannosauroida in morphospace occupation (Figs 2C, 6C).

Some authors have advocated caution in using discrete character cladistic data matrices as a source of data for disparity studies (Anderson & Friedman 2012; Benson 2018). These authors note several advantages of such data sets: they are readily available, they typically document broad anatomical coverage, they enable comparisons between taxa of very different form, ancestral taxa can be reconstructed, and they can be scored for fragmentary material. On the other hand, the use of cladistic data matrices entails some problems: it is unclear how to relate anatomical and functional variation both for individual characters or for overall morphologies; such data matrices may reflect phylogenetic signal rather than any aspect of ecomorphology; they may concentrate on obscure anatomical details of, say, the braincase or maxilla orientation, which may have limited functional or evolutionary importance; other significant characters are ignored because they are either autapomorphies or record shape but not phylogenetic signal. Benson (2018) gives examples of published data matrices that may give apparently high rates for certain clades, such as birds, but it cannot be said whether the evolutionary rates are truly high or simply reflect excessive research interest and coding of many small-scale anatomical characters. Gerber (2019) highlights problems of excessive missing data and the meaning of the axes in cases where cladistic characters are used, but they and others (Lloyd 2016; Hopkins & Gerber 2017) support the use of discrete characters to document disparity when appropriate methods are used to minimize bias and understand the data. Equally, of course, outline and landmark studies may be capturing aspects of shape that have no evolutionary or ecological significance at all or might be simply size-related (Gould 1966).

Anderson & Friedman (2012) recommend that analysts devise metrics that have functional significance, such as ratios of relative lengths of portions of the jaw or limbs. Another productive approach could be to use subsets of discrete character data sets that have hypothesized ecomorphological or evolutionary significance (Gerber 2019; Stubbs *et al.* 2019), such as the appendicular skeleton in the fish–tetrapod transition (Ruta & Wills 2016), the display crests and feeding apparatus in hadrosaurid dinosaurs (Stubbs *et al.* 2019) and this contribution. Alternatively, researchers could construct or combine new matrices from multiple sources for disparity studies that

focus on specific parts of anatomy, such as the teeth (Strickson *et al.* 2016) or teeth and jaw (Nordén *et al.* 2018).

Ecomorphology

Several ecomorphological features can be seen in our results. First, it is very noticeable that oviraptorids dominate all morphospace analyses (Figs 2, 6). Their bizarre mandible shape dominates PC1 and separates them from the general cluster of taxa. Oviraptorids were also identified as an aberrant group in the studies of skull shape by Brusatte *et al.* (2012) and Foth & Rauhut (2013). Second, the maniraptoriforms have an overall higher disparity than the tyrannosauroids. This reflects the high disparity of the oviraptorids, but also the fact that the clade includes birds and their ancestors, which were mainly carnivores and insectivores, as well as omnivores and herbivores, some without teeth. These maniraptoriforms encompassed a substantially greater amount of total phylogenetic branch duration than tyrannosauroids, giving them more opportunity to accumulate ecological diversity.

It is interesting that the omnivore–herbivore category, even though sparsely sampled, occupies a great area of morphospace according to all methods (Figs 3, 7) and this, as noted earlier, may simply reflect the fact that this is an ill-defined or waste-basket dietary category. The relative amount of morphospace is especially pronounced with the shape methods (cf. Figs 3D–E, 7D–E). This too reflects the high disparity of oviraptorids, and the generally astonishing variety of shapes of mandibles in that clade, the ornithomimids and others. The differences are less pronounced in the discrete character analyses (Figs 3F, 7F) perhaps because relatively few cladistic-style characters are used to describe ecomorphological differences in the mandible.

CONCLUSION

Our results suggest that all methods of measuring morphological disparity give comparable results, and this is not solely the result of phylogeny. In our case, the two shape-based methods (outlines, landmarks) gave very similar results as they are both recording the same features and rendering them on the dominant multivariate axes in the same ways. We did note some subtle differences in recording of jaw attenuation and bending between the methods, and we highlight the importance of exploring shape changes along axes beyond those shown in biplots of PC1 and PC2 morphospace. By comparing different sampling regimes our comparisons show that a

reduced sample of taxa recover the same major axes of shape variation and a full sample with more taxa largely saturates morphospace without considerably expanding it. Our key findings, that the disparity of the major clades and dietary categories were comparable among all three methods, matches earlier findings from other such comparisons of disparity data sets for different taxa. But, as we noted, each method gives subtly different degrees of separation and overlap in disparity between different clades, and so we recommend that future studies should use multiple approaches when assessing disparity as each has its advantages, and each data type reflects different aspects of morphology in relation to function and evolution.

Acknowledgements. We thank members of the Bristol Palaeobiology group for feedback and Mark Puttick, Vincent Bonhomme, Marcello Ruta, and Suresh Singh for their methodological advice. We also thank two anonymous reviewers for their feedback which improved this manuscript. This MS comes from an MSc thesis submitted by JS as part of the requirements of the MSc in Palaeobiology at the University of Bristol. Funding to MJB and TJS from UK Natural Environment Research Council grant (NE/P013724/1) and ERC Advanced Grant (788203 INNOVATION).

Author contributions. TLS, EJR and MJB designed and supervised the study, TLS wrote code and developed the methods, JS and TLS carried out the analyses and wrote the first version of the MS, and all authors contributed to the final version.

DATA ARCHIVING STATEMENT

Data for this study are available in the Dryad Digital Repository: <https://doi.org/10.5061/dryad.70rk2cf>

Editor. Kenneth Angielczyk

REFERENCES

- ADAMS, D. C. and OTÁROLA-CASTILLO, E. 2013. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology & Evolution*, **4**, 393–399.
- ANDERSON, P. S. L. 2008. Shape variation between arthrodire morphotypes indicates possible feeding niches. *Journal of Vertebrate Paleontology*, **28**, 961–969.
- and FRIEDMAN, M. 2012. Using cladistic characters to predict functional variety: experiments using early gnathostomes. *Journal of Vertebrate Paleontology*, **32**, 1254–1270.
- BRAZEAU, M. D. and RAYFIELD, E. J. 2011. Initial radiation of jaws demonstrated stability despite faunal and environmental change. *Nature*, **476**, 206–209.
- BAPST, D. W. 2012. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology & Evolution*, **3**, 803–807.

- BULLOCK, P. C., MELCHIN, M. J., SHEETS, H. D. and MITCHELL, C. E. 2012. Graptoloid diversity and disparity became decoupled during the Ordovician mass extinction. *Proceedings of the National Academy of Sciences*, **109**, 3428–3433.
- BENSON, R. J. 2018. Dinosaur macroevolution and macroecology. *Annual Review of Ecology, Evolution, & Systematics*, **49**, 379–408.
- BONHOMME, V., PICQ, S., GAUCHEREL, C. and CLAUDE, J. 2014. Momocs: outline analysis using R. *Journal of Statistical Software*, **56**, 1–24.
- BOOKSTEIN, F. L. 1991. *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, 456 pp.
- BRUSATTE, S. L., BENTON, M. J., RUTA, M. and LLOYD, G. T. 2008. The first 50 Myr of dinosaur evolution: macroevolutionary pattern and morphological disparity. *Biology Letters*, **4**, 733–736.
- SAKAMOTO, M., MONTANARI, S. and HARCOURT SMITH, W. E. H. 2012. The evolution of cranial form and function in theropod dinosaurs: insights from geometric morphometrics. *Journal of Evolutionary Biology*, **25**, 365–377.
- LLOYD, G. T., WANG, S. C. and NORELL, M. A. 2014. Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur–bird transition. *Current Biology*, **24**, 2386–2392.
- CABRERA, J. M. and GIRI, F. 2016. Evomorph: evolutionary morphometric simulation. <https://cran.r-project.org/web/packages/Evomorph/index.html>
- CAILLIEZ, F. 1983. The analytical solution of the constant problem. *Psychometrika*, **48**, 305–308.
- CARDINI, A. 2016. Lost in the other half: improving accuracy in geometric morphometric analyses of one side of bilaterally symmetric structures. *Systematic Biology*, **65**, 1096–1106.
- CIAMPAGLIO, C. N., KEMP, M. and McSHEA, D. W. 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology*, **27**, 695–715.
- CLOSE, R. A., FRIEDMAN, M., LLOYD, G. T. and BENSON, R. B. J. 2015. Evidence for a mid-Jurassic adaptive radiation in mammals. *Current Biology*, **25**, 2137–2142.
- COONEY, C. R., BRIGHT, J. A., CAPP, E. J. R., CHIRA, A. M., HUGHES, E. C., MOODY, C. J. A., NOURI, L. O., VARLEY, Z. K. and THOMAS, G. H. 2017. Mega-evolutionary dynamics of the adaptive radiation of birds. *Nature*, **542**, 344–347.
- DELINE, B., GREENWOOD, J. M., CLARK, J. W., PUTTICK, M. N., PETERSON, K. J. and DONOGHUE, P. C. J. 2018. Evolution of metazoan morphological disparity. *Proceedings of the National Academy of Sciences*, **115**, 8909–8918.
- ERWIN, D. H. 2007. Disparity: morphological patterns and developmental context. *Palaentology*, **50**, 57–73.
- FELSENSTEIN, J. 1985. Phylogenies and the comparative method. *The American Naturalist*, **125**, 1–15.
- FOOTE, M. 1997. Evolution of morphological diversity. *Annual Review of Ecology & Systematics*, **28**, 129–152.
- FORTEY, R. A., BRIGGS, D. E. G. and WILLS, M. A. 1996. The Cambrian evolutionary ‘explosion’: decoupling cladogenesis from morphological disparity. *Biological Journal of the Linnean Society*, **57**, 13–33.
- FOTH, C. and RAUHUT, O. W. 2013. Macroevolutionary and morphofunctional patterns in theropod skulls: a morphometric approach. *Acta Palaeontologica Polonica*, **58**, 1–16.
- BRUSATTE, S. L. and BUTLER, R. J. 2012. Do different disparity proxies converge on a common signal? Insights from the cranial morphometrics and evolutionary history of Pterosauria (Diapsida: Archosauria). *Journal of Evolutionary Biology*, **25**, 904–915.
- FRECKLETON, R. P., HARVEY, P. H. and PAGEL, M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, **160**, 712–726.
- FRIEDMAN, M. 2009. Ecomorphological selectivity among marine teleost fishes during the end-Cretaceous extinction. *Proceedings of the National Academy of Sciences*, **106**, 5218–5223.
- GERBER, S. 2019. Use and misuse of discrete character data for morphospace and disparity analyses. *Palaentology*, **62**, 305–319.
- GOSWAMI, A. and POLLY, P. D. 2010. The influence of modularity on cranial morphological disparity in Carnivora and Primates (Mammalia). *PLoS One*, **5** (3), e9517.
- GOULD, S. J. 1966. Allometry and size in ontogeny and phylogeny. *Biological Reviews*, **41**, 587–638.
- GROSSNICKLE, D. M. and POLLY, P. D. 2013. Mammal disparity decreases during the Cretaceous angiosperm radiation. *Proceedings of the Royal Society B*, **280**, 20132110.
- GUILLERME, T. 2018. dispRity: a modular R package for measuring disparity. *Methods in Ecology & Evolution*, **9**, 1755–1763.
- HARVEY, P. H. and PAGEL, M. D. 1991. *The comparative method in evolutionary biology*. Oxford University Press, 248 pp.
- HETHERINGTON, A. J., SHERRATT, E., RUTA, M., WILKINSON, M., DELINE, B. and DONOGHUE, P. C. J. 2015. Do cladistic and morphometric data capture common patterns of morphological disparity? *Palaentology*, **58**, 393–399.
- HILL, J. J., PUTTICK, M. N., STUBBS, T. L., RAYFIELD, E. J. and DONOGHUE, P. C. J. 2018. Evolution of jaw disparity in fishes. *Palaentology*, **61**, 847–854.
- HOPKINS, M. J. 2017. How well does a part represent the whole? A comparison of cranial shape evolution with exoskeletal character evolution in the trilobite family Pteroccephaliidae. *Palaentology*, **60**, 309–318.
- and GERBER, S. 2017. Morphological disparity. 1–12. In NUÑO DE LA ROSA, L. and MÜLLER, G. B. (eds). *Evolutionary developmental biology*. Springer, 2000 pp.
- HUGHES, M., GERBER, S. and WILLS, M. A. 2013. Clades reach highest morphological disparity early in their evolution. *Proceedings of the National Academy of Sciences*, **110**, 13875–13879.
- LAURIN, M. 2004. The evolution of body size, Cope’s rule and the origin of amniotes. *Systematic Biology*, **53**, 594–622.
- LLOYD, G. T. 2016. Estimating morphological diversity and tempo with discrete character–taxon matrices:

- implementation, challenges, progress, and future directions. *Biological Journal of the Linnean Society*, **118**, 131–151.
- MACLAREN, J. A., ANDERSON, P. S. L., BARRETT, P. M. and RAYFIELD, E. J. 2017. Herbivorous dinosaur disparity and its relationship to extrinsic evolutionary drivers. *Paleobiology*, **43**, 15–33.
- MACLEOD, N. 1999. Generalizing and extending the eigen-shape method of shape space visualization and analysis. *Paleobiology*, **25**, 107–138.
- MITTEROECKER, P. and GUNZ, P. 2009. Advances in geometric morphometrics. *Evolutionary Biology*, **36**, 235–247.
- MONGIARDINO KOCH, N., CECCARELLI, F. S., OJANGUREN-AFFILASTRO, A. A. and RAMÍREZ, M. J. 2017. Discrete and morphometric traits reveal contrasting patterns and processes in the macroevolutionary history of a clade of scorpions. *Journal of Evolutionary Biology*, **30**, 814–825.
- MONTEIRO, L. R. and NOGUEIRA, M. R. 2011. Evolutionary patterns and processes in the radiation of phyllostomid bats. *BMC Evolutionary Biology*, **11**, 137.
- NAVARRO, C. A., MARTIN-SILVERSTONE, E. and STUBBS, T. L. 2018. Morphometric assessment of pterosaur jaw disparity. *Royal Society Open Science*, **5**, 172130.
- NORDÉN, K. K., STUBBS, T. L., PRIETO-MÁRQUEZ, A. and BENTON, M. J. 2018. Multifaceted disparity approach reveals dinosaur herbivory flourished before the end-Cretaceous mass extinction. *Paleobiology*, **44**, 620–637.
- OKSANEN, A. J., BLANCHET, F. G., KINDT, R., LEGENDRE, P., MINCHIN, P. R., HARA, R. B. O., SIMPSON, G. L., SOLYMOS, P., STEVENS, M. H. H. and WAGNER, H. 2017. vegan: Community Ecology Package. R package version 2.4-3. <https://cran.r-project.org/web/packages/vegan/index.html>
- PARADIS, E. and SCHLIEP, K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, **35**, 526–528.
- POLLY, P. D. 2004. On the simulation of the evolution of morphological shape: multivariate shape under selection and drift. *Palaeontologia Electronica*, **7**, 1–28.
- R CORE TEAM. 2017. R: a language and environment for statistical computation. v. 3.4.3. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>
- REVELL, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology & Evolution*, **3**, 217–223.
- ROHLF, F. J. 2017a. tpsUtil, v. 1.74. Department of Ecology & Evolution, State University of New York, Stony Brook. <https://life.bio.sunysb.edu/morph/>
- 2017b. TpsDig, v. 2.30. Department of Ecology & Evolution, State University of New York, Stony Brook.
- 2017c. tpsRelw, v. 1.68. Department of Ecology & Evolution, State University of New York, Stony Brook. <https://life.bio.sunysb.edu/morph/>
- ROMANO, M. 2017. Disparity vs. diversity in Stegosauria (Dinosauria, Ornithischia): cranial and post-cranial sub-dataset provide different signals. *Historical Biology*, **31**, 857–865.
- RUTA, M. and WILLS, M. A. 2016. Comparable disparity in the appendicular skeleton across the fish–tetrapod transition, and the morphological gap between fish and tetrapod postcrania. *Palaeontology*, **59**, 249–267.
- ANGIELCZYK, K. D., FROBISCH, J. and BENTON, M. J. 2013. Decoupling of morphological disparity and taxic diversity during the adaptive radiation of anomodont therapsids. *Proceedings of the Royal Society B*, **280**, 20131071.
- SCHAEFFER, J., BENTON, M. J., RAYFIELD, E. J. and STUBBS, T. J. 2019. Data from: morphological disparity in theropod jaws and the application of discrete characters and geometric morphometrics. *Dryad Digital Repository*. <https://doi.org/10.5061/dryad.70rk2cf>
- SMITHWICK, M. and STUBBS, T. L. 2018. Phanerozoic survivors: actinopterygian evolution through the Permo-Triassic and Triassic-Jurassic mass extinction events. *Evolution*, **72**, 348–362.
- STRICKSON, E., PRIETO-MÁRQUEZ, A., BENTON, M. J. and STUBBS, T. L. 2016. Dynamics of dental evolution in ornithomimid dinosaurs. *Scientific Reports*, **6**, 28904.
- STUBBS, T. L., PIERCE, S. E., RAYFIELD, E. J. and ANDERSON, P. S. L. 2013. Morphological and biomechanical disparity of crocodile-line archosaurs following the end-Triassic extinction. *Proceedings of the Royal Society B*, **280**, 20131940.
- BENTON, M. J., ELSLER, A. and PRIETO-MÁRQUEZ, A. 2019. Morphological innovation and the evolution of hadrosaurid dinosaurs. *Paleobiology*, **45**, 347–362.
- VILLIER, L. and EBLE, G. J. 2004. Assessing the robustness of disparity estimates: the impact of morphometric scheme, temporal scale, and taxonomic level in spatangoid echinoids. *Paleobiology*, **30**, 652–665.
- WATANABE, A. 2018. How many landmarks are enough to characterize shape and size variation? *PLoS One*, **13**, e0198341.
- WEISHAMPEL, D., DODSON, P. and OSMÓLSKA, H. 2004. *The Dinosauria*. 2nd edn. University of California Press, 880 pp.
- WILLS, M. A., BRIGGS, D. E. G. and FORTEY, R. A. 1994. Disparity as an evolutionary index: a comparison of Cambrian and recent arthropods. *Paleobiology*, **20**, 93–130.
- ZANNO, L. E. and MAKOVICKY, P. J. 2011. Herbivorous ecomorphology and specialization patterns in theropod dinosaur evolution. *Proceedings of the National Academy of Sciences*, **108**, 232–237.
- ZELDITCH, M. L., SWIDERSKI, D. L. and SHEETS, H. D. 2012. *Geometric morphometrics for biologists: a primer*. 2nd edn. Academic Press, 488 pp.
- LI, J., TRAN, L. A. and SWIDERSKI, D. L. 2015. Relationships of diversity, disparity, and their evolutionary rates in squirrels (Sciuridae). *Evolution*, **69**, 1284–1300.